ABSTRACT OF THE DISCLOSURE

A three-dimensional structure estimation apparatus by which an image of a wide visual field and another image which has a high resolution at part of the visual field are provided and a high resolution three-dimensional structure of an object in the visual field can be estimated is disclosed. A wide visual field image and a narrow visual field image outputted from cameras which have a wide visual field cone and a narrow visual field cone of different visual fields and produce images of different resolutions are converted into images whose pixel units are equal in magnitude by a conversion section. The conversion section includes a sampling section which samples pixels of the narrow visual field image to produce a coincident pixel unit image which has a pixel unit coincident with that of the wide visual field image, and a depth image production section which receives and compares the wide visual field image and the coincident pixel unit image with each other to produce a depth image. Preferably, the sampling section produces a plurality of images of coincident pixels by successively displacing sampling positions little by little so that the depth distance can be estimated with a higher degree of accuracy.